AGU100 ADVANCING EARTH AND SPACE SCIENCE

Geophysical Research Letters

RESEARCH LETTER
10.1029/2018GL081686

Key Points:
- A novel method for climate model evaluation is used to identify both mean errors and potential compensating errors
- As an example, we apply this methodology to investigate the quality of different cloud parameterizations over the Southern Ocean
- Changes to the cloud phase parameterizations can reduce shortwave radiative bias regionally, but large compensating errors remain

Supporting Information:
- Supporting Information S1

Correspondence to:
A. J. Schuddeboom,
alex.schuddeboom@pg.canterbury.ac.nz

Citation:
Schuddeboom, A., Varma, V., McDonald, A. J., Morgenstern, O., Harvey, M., Parsons, S., et al. (2019). Cluster-based evaluation of model compensating errors: A case study of cloud radiative effect in the Southern Ocean. *Geophysical Research Letters*, *46*, 3446–3453. https://doi.org/10.1029/2018GL081686

Received 13 DEC 2018
Accepted 4 MAR 2019
Accepted article online 11 MAR 2019
Published online 28 MAR 2019

Cluster-Based Evaluation of Model Compensating Errors: A Case Study of Cloud Radiative Effect in the Southern Ocean

Alex Schuddeboom[1], Vidya Varma[2], Adrian J. McDonald[1], Olaf Morgenstern[2], Mike Harvey[2], Simon Parsons[1], Paul Field[3], and Kalli Furtado[3]

[1]School of Physical and Chemical Sciences, University of Canterbury, Christchurch, New Zealand, [2]National Institute of Water and Atmospheric Research, Wellington, New Zealand, [3]MetOffice, Exeter, UK

Abstract Model evaluation is difficult and generally relies on analysis that can mask compensating errors. This paper defines new metrics, using clusters generated from a machine learning algorithm, to estimate mean and compensating errors in different model runs. As a test case, we investigate the Southern Ocean shortwave radiative bias using clusters derived by applying self-organizing maps to satellite data. In particular, the effects of changing cloud phase parameterizations in the MetOffice Unified Model are examined. Differences in cluster properties show that the regional radiative biases are substantially different than the global bias, with two distinct regions identified within the Southern Ocean, each with a different signed bias. Changing cloud phase parameterizations can reduce errors at higher latitudes but increase errors at lower latitudes of the Southern Ocean. Ranking the parameterizations often shows a contrast in mean and compensating errors, notably in all cases large compensating errors remain.

1. Introduction

Climate models are validated with comparisons to observational data; however, many of these comparisons are simplistic, focusing extensively on mean or zonally averaged behavior. These approaches can hide compensating errors, giving the false appearance of accuracy (Hyder et al., 2018; Jakob, 2003). The goal of this paper is to provide an approach for the evaluation of a model based on both its average error and its compensating errors. This technique presents a more nuanced approach to model evaluation which avoids the aforementioned problems.

One of the longest standing problems in climate models is the Southern Ocean radiative bias (Wild et al., 1995), in which a weak shortwave cloud radiative effect (CRE) over the Southern Ocean leads to an excess in surface downwelling solar radiation (Bodas-Salcedo et al., 2012; Kay et al., 2012). The most common explanation for this phenomena is that the models do not adequately capture cloud phase, with particular issues accurately representing supercooled liquid water (Bodas-Salcedo et al., 2014, 2016). However, other factors such as a lack of low-level cloud cover are also likely contributing factors (Bodas-Salcedo et al., 2012; Schuddeboom et al., 2018). Previous studies have used several different approaches to attempt to solve the cloud phase issue. Examples include Kay et al. (2016) who analyzed the impact of changing the ice nucleation temperature parameterization, Furtado and Field (2017) who examined the effects of changing several ice process parameterizations and Furtado et al. (2016) who showed that a physical-based method for diagnosing mixed-phase cloud fractions (CFs) could significantly reduced net radiation errors. Given the importance of ice processes to clouds within the Southern Ocean region, this work focuses on investigating the impacts of altering parameterizations related to ice formation in the model. This analysis is primarily used as an exploration of the model evaluation methodology, with investigating model issues acting as a secondary objective which will be the subject of a later paper.

The cloud clusters developed in Schuddeboom et al. (2018) are used to account for the effect of compensating errors. These clusters were generated by applying a self-organizing map (SOM; Kohonen, 1998, 2013) to Moderate Resolution Imaging Spectroradiometer (MODIS) cloud data. Several previous papers have also used the cluster framework to analyze model performance including Williams and Tselioudis (2007), Williams and Webb (2009), Mason et al. (2015), Jin et al. (2016), Jin et al. (2017), and Tan et al. (2017). The

Table 1
The Observational Values for Relative Frequency of Occurrence (RFO), Shortwave Cloud Radiative Effect (SW CRE), Cloud Fraction (CF), Cloud Top Pressure (CTP), and Cloud Optical Thickness (COT) for Each of the Clusters

Cluster #	1	2	3	4	5	6	7	8	9	10	11	12
Regime	Tr	Tr	Mr	Mr	LB	LB	Mr	Mr	ML	ML	St	St
RFO (%)	12.1	4.8	10.9	13.7	5.1	13.1	3.9	4.6	13.4	5.4	3.6	9.4
SW CRE (W/m^2)	−50	−18	−19	−33	−65	−14	−34	−45	−104	−84	−57	−75
CF (%)	78	35	50	71	75	26	55	74	93	91	75	89
CTP (hPa)	280	680	800	860	450	630	760	830	350	680	750	850
COT (τ)	4.5	6.1	3.6	5.2	18.2	19.7	10.3	9.5	20.9	26.5	20.2	16.1

Note. The corresponding abbreviations for each of the regimes are Tr for tropical cloud, Mr for marine cloud, LB for land-based cloud, ML for mixed layer, and St for stratocumulus.

development of the global cloud regime error metric for the present day ($CREM_{pd}$) described in Williams and Webb (2009) is a particularly important influence for the methods described here.

2. Data and Methods

2.1. Observational Data

The observational data sets used in this analysis are the MODIS Collection 6 data set (Platnick et al., 2003, 2017) and the Clouds and the Earth's Radiant Energy System (CERES) synoptic 1° (SYN1deg) Edition 3A data set (Wielicki et al., 1996). In particular, we use cloud top pressure-cloud optical thickness (CTP-COT) joint histograms from the MODIS data set and daily upward radiative fluxes from the CERES data set. The MODIS CTP-COT histograms were used to generate the clusters in Schuddeboom et al. (2018), which are reused in this analysis. The clusters are included in supporting information as Figure S1 along with additional text describing their characteristics. The average values of several key variables for each of these clusters are shown in Table 1. Additionally, each of the clusters was assigned to a regime in Schuddeboom et al. (2018) that is also included in Table 1. The MODIS data are used to determine the CTP, COT, relative frequency of occurrence (RFO), and CF of each cluster. The CERES data are used to calculate the shortwave CRE (SW CRE). Detailed descriptions of the data preparation and clustering processes used here are given in McDonald et al. (2016) and Schuddeboom et al. (2018).

2.2. Model Data

The climate model used in this study is the free running Global Atmosphere (GA) version 7.1 of the MetOffice Unified Model (Brown et al., 2012; Walters et al., 2017). The resolution of the simulation is N96 (i.e., 1.875° ×1.25°) in the horizontal and 85 model levels in vertical. The model uses the ENDGame dynamical core with a semiimplicit semi-Lagrangian formulation to solve the nonhydrostatic, fully compressible deep-atmosphere equations of motion Wood et al. (2014). On top of the standard GA7.1 version described in Walters et al. (2017), a modified microphysics scheme is used. This includes some changes to the riming process, such as a shape dependence of riming rates using the parametrization by Heymsfield and Miloshevich (2003), the prevention of small liquid droplets from riming and some convection-permitting modeling (Furtado & Field, 2017). In the present study, the GA7.1 control run follows the free running Atmospheric Model Intercomparison Project climate model development protocol (Gates et al., 1999). To enable clustering with the model data, equivalent MODIS CTP-COT histograms are generated from the model using the COSP simulator (Bodas-Salcedo et al., 2011). These histograms are then assigned to one of the MODIS clusters based on the Euclidean distance between the histogram and the clusters, as described in Schuddeboom et al. (2018).

As excess of ice is a known issue in the atmospheric model configurations of the Unified Model, our sensitivity tunings in this study are focused on the microphysics scheme that controls ice formation. After testing several combinations of ice formation processes in the model, deposition seems to be the most influential. In this study, two different approaches are adopted that incorporate changes to the ice deposition scheme. The first is tuning the values of the ice deposition rate. In the cloud microphysics scheme, the ice deposition rate is multiplied by a factor to make the growth and sublimation processes asymmetric. The default value used is 0.9. For our sensitivity runs, we use three different values for this factor 0, 0.01, and 0.6. Changing this parameter reduces the ice growth but does not effect the sublimation. As mentioned earlier, the primary

purpose of this study is to investigate the validity of the approaches used rather than testing potential model changes; hence, it does not matter that modifications to this deposition factor are unphysical. The second parameter that was selected for tuning in the microphysics scheme is the shape parameter or the capacitance of the ice crystals. The growth rate of ice crystals is given by the equation,

$$\frac{dm}{dt} = 4\pi CD(\rho v - \rho v r), \tag{1}$$

where C is the capacitance of the ice crystals. Furtado et al. (2015) showed that the relative humidity in ice clouds can be strongly effected by the choice of C, indicating that it could have strong effect on liquid water contents under mixed-phase conditions. By default, the model assumes that all ice crystals are spherical and hence uses the default value of $C = 1$. Values of 0.25, 0.5, and 0.75 for C are used to test the model sensitivity to this parameter. A further set of changes are introduced in addition to the modified capacitance rate to form the "capacitance +" runs. In those runs the ice nucleation temperature, the temperature at which heterogeneous nucleation of ice first starts to form in the model is changed from the default value of −10 to −40 °C. Two additional parameters that control the ice formation in the model are also changed. These are the temperature that defines the rate at which condensate gets detrained as ice and the temperature at which all condensate is detrained as ice which have been modified to −40 and −41 °C, respectively, from the original values of −10 and −20 °C.

2.3. Methodology

The most straightforward way to evaluate the quality of the model using clusters is to directly compare the mean behavior of each cluster. These comparisons are focused on the mean RFO and CRE as these two variables capture the radiative behavior of cloud. More detailed information can be gathered from this approach by restricting the analysis to particular subregions of the globe. While this provides detail about the model, it can also makes it difficult to judge the relative performance of different model runs due to the sheer volume of information. To make these comparisons simpler and more quantitative, we use the following equations to calculate the magnitude of CRE model errors:

$$\Delta\text{CRE} = |\sum_N C_N^{\text{Model}} R_N^{\text{Model}} - C_N^{\text{CERES}} R_N^{\text{MODIS}}|, \tag{2}$$

$$|\text{CRE}| = \sum_N |C_N^{\text{Model}} R_N^{\text{Model}} - C_N^{\text{CERES}} R_N^{\text{MODIS}}|, \tag{3}$$

where C represents CRE corresponding to a cluster N and R represents the RFO of cluster N. While similar, ΔCRE and |CRE| serve a different purpose. ΔCRE is the magnitude of the mean difference in CRE between the model and observations, while |CRE| is the integrated cluster error. This means that |CRE| represents the model error when compensating errors between clusters are not able to cancel each other out. While this approach does not account for the compensating errors that occur within each of the clusters, it should represent the bulk of the compensating errors in the model runs. The |CRE| is similar to the $CREM_{\text{pd}}$ in Williams and Webb (2009) and the *TAFB* in Hyder et al. (2018). While the $CREM_{\text{pd}}$ could be used in place of the |CRE|, the values it generates are not as easily comparable to the mean error values. We choose to use the |CRE| as it is simpler and provides a more physically interpretable value. Directly including ΔCRE in the evaluation process is also an improvement over Williams and Webb (2009) as it better contextualizes the |CRE| results. This approach could also be extended to other properties such as CF, cloud phase, or precipitation. Like the averaging approach mentioned above, ΔCRE and |CRE| can be used to provide extra information by restricting the analysis to specific regions.

3. Results and Discussion

The average values of RFO and SW CRE describe the behavior of a given cluster. To compare the clusters in each of the model realizations, while keeping the observations in focus, these variables are displayed as anomalies from the observational values for each cluster in Figure 1. This figure shows the RFO and SW CRE for each cluster both globally and over the Southern Ocean (defined here as between 40°S and 70°S with data over land that is within this band also included). It is immediately clear that the model runs agree more strongly with each other than the observations. The global SW CRE shows clear signs of potential compensating errors with a balance of positive and negative biased clusters, whereas the Southern Ocean consists of mostly negative biased clusters.

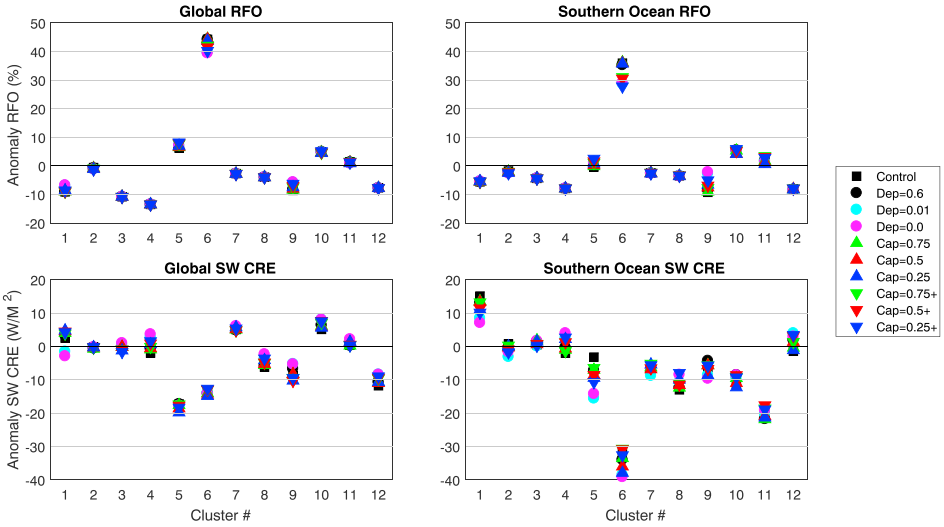

Figure 1. The anomaly in mean relative frequency of occurrence (RFO) and shortwave cloud radiative effect (SW CRE) between the observational data and each of the model runs for all of the clusters. Values are shown both over the globe and over the Southern Ocean. Here the Southern Ocean is defined as within the boundaries of 40°S and 70°S. The corresponding global observational values are shown in Table 1.

The most notable cluster in Figure 1 is cluster 6 as it shows the largest discrepancies in RFO, both globally and over the Southern Ocean, as well as relatively large anomalies in SW CRE. This cluster also shows substantial variance between the different model realizations. Other notable clusters include clusters 3 and 4, which show decreases in occurrence rates in all the model runs, and clusters 5 and 9, which show large differences in SW CRE both globally and over the Southern Ocean. The cluster biases in RFO over the globe and the Southern Ocean show strong agreement. However, the SW CRE biases show significant differences between the global and Southern Ocean regions, particularly with clusters 5, 6, 7, 10, 11, and 12. Clusters 10–12 were also identified as particularly important in the Southern Ocean region in Schuddeboom et al. (2018). The wider range of SW CRE values in the runs over the Southern Ocean demonstrates how important ice parameterizations are in this region. Additionally, particular clusters, such as clusters 5 and 6, are shown to be especially responsive to changes in the ice formation parameterizations. The corresponding longwave CRE (LW CRE) results are included as supporting information Figure S2 and also identify substantial variations in clusters 5 and 6.

Focusing directly on the Southern Ocean, Figure 2 shows the difference in SW CRE between the model runs and the CERES data over the Southern Ocean. The distribution of SW CRE displays two distinct regional biases: a negative bias in the lower latitudes of the Southern Ocean and a positive bias in the higher latitudes. It is clear that any analysis that averages these two regions together would mask very large compensating errors. These two regions are referred to as Top Southern Ocean (40°S to 55°S) and Bottom Southern Ocean (55°S to 70°S) for the remainder of this paper. The borders of these regions are indicated as black lines on Figure 2 and show a clean separation between the regions. The LW CRE results, included as supporting information Figure S3, are more geographically variable; however, the same partitioning appears valid. While the SW CRE is similar in all of the model realizations, the model runs with a stronger Top Southern Ocean bias (such as the $Dep = 0.01$, $Dep = 0$, and $Cap = 0.25$ runs) tend to show a weaker Bottom Southern Ocean bias. The opposite effect is also visible in model runs with a strong Bottom Southern Ocean bias (such as the Control, $Cap = 0.75+$ and $Dep = 0.6$ runs) showing a weaker Top Southern Ocean bias. This suggests that the biases in one region are inherently linked to the biases in the other. A likely explanation is that changing the parameterization strongly effects the brightness of clouds due to greater liquid water contents, leading to a uniform response over the Southern Ocean. The model runs also show a greater range of results in the Bottom Southern Ocean than the Top Southern Ocean. Interestingly, there appears to be little to no link between how a model run performs in SW CRE and LW CRE, which suggests that these changes impact more than just the phase of clouds.

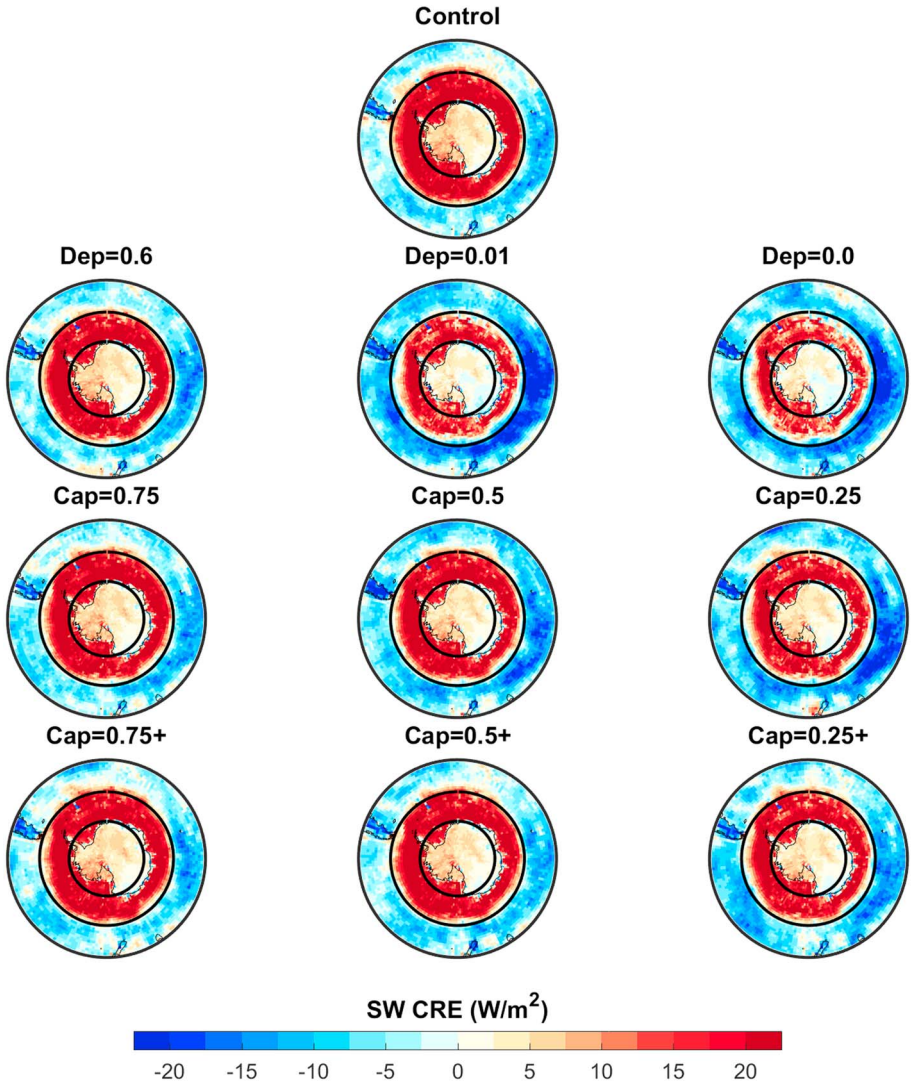

Figure 2. The difference between the shortwave cloud radiative effect (SW CRE) in the model runs and the Clouds and the Earth's Radiant Energy System data set over the Southern Ocean region.

While the earlier results demonstrate clear differences between the model runs and the observations, judging the relative accuracy, each of the model runs is still difficult. For this reason, the equations (2) and (3) are used to generate Figure 3. This figure allows for the rapid evaluation of the performance of the different model runs, with the x axis corresponding to ΔCRE and the y axis to |CRE|. The simplest way to interpret Figure 3 is that the closer a model run is to the origin of the axis, the better it agrees with the observations. It is also important to consider that while this paper is focused on reducing the compensating errors, achieving this at the expense of the mean error is not productive. Ideally, compensating errors would be addressed by removing all the errors that are contributing as only removing one of these errors can adversely effect simulation quality.

Several features can be immediately identified by the comparison of the model runs in Figure 3. The two low deposition rate runs are the best performing model runs for global SW CRE in both mean and compensating errors. However, when constrained to the Southern Ocean, they may still perform well with respect to compensating errors but are the worst performing model runs in mean error. Although the impact of changing these parameterizations appears to result in a global improvement, it makes the mean problems over the Southern Ocean worse. The fact that they still perform well in compensating errors over the Southern Ocean suggests that the deposition changes could potentially be paired with other changes to lead to true model improvement. When both mean and compensating errors are considered, the capacitance + runs appear

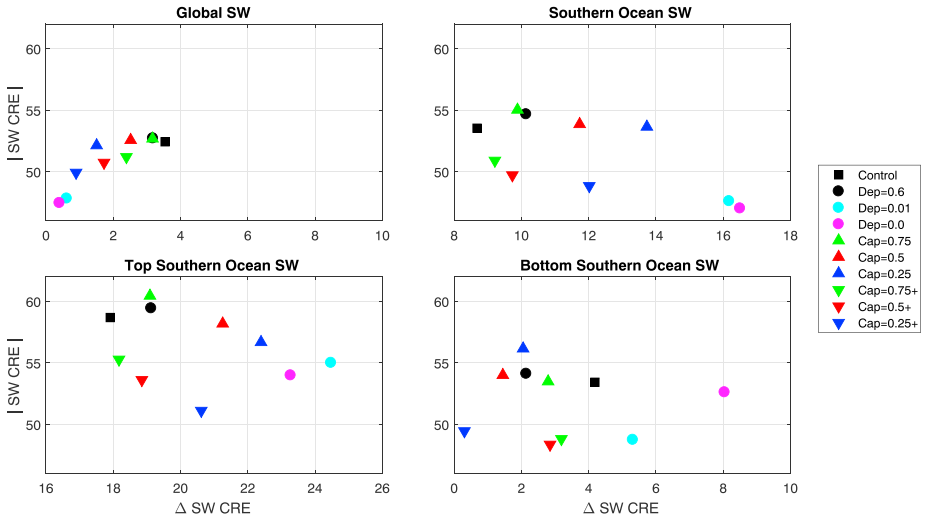

Figure 3. The ΔSW CRE and |SW CRE| for each of the models. The upper left subplot shows the global average values while the others correspond to subregions of the Southern Ocean. The upper right subplot covers the entire Southern Ocean region, defined here as between 40°S and 70°S. The bottom row shows comparable plots restricted to the Top Southern Ocean and Bottom Southern Ocean regions. The Top Southern Ocean region is the region from 40°S to 55°S, and the Bottom Southern Ocean is from 55°S to 70°S. The variables on each of the axis are described in equations (2) and (3). SW CRE = shortwave cloud radiative effect.

to be the best performers over the Southern Ocean region. Additionally, the capacitance + runs appear to consistently outperform their regular capacitance equivalents in compensating errors. This suggests that it would be worth testing these additional modifications independently of the capacitance changes as they may reduce model compensating errors in general. Several model runs also appear to exhibit drastically different behaviors in the Top Southern Ocean and Bottom Southern Ocean regions. The LW CRE equivalent figure is included in the supporting information as Figure S4 and shows a radically different model ranking. This suggests that while some of the changes to the model parameterizations may improve SW CRE, they could adversely effect the LW CRE.

Examining the magnitude of the biases in Figure 3, we see that the size of the compensating errors is much larger than the mean errors. For example, the global panel shows mean errors around 2 W/m^2 and a total compensating error of around 50 W/m^2 suggesting a 26-W/m^2 error that is mostly cancelled out by a 24-W/m^{2}error. In the Southern Ocean and Top Southern Ocean regions the mean biases are much larger and the total compensating error is around the same magnitude. When combined with the result that almost all of the clusters in Figure 1 are biased toward a more negative SW CRE over the Southern Ocean, this indicates that compensating errors between cloud types are less dominant over the Southern Ocean than they are globally. Instead, the Southern Ocean is dominated by a more consistent bias that effects the majority of the clusters. For example, the Top Southern Ocean region shows mean errors around 20 W/m^2 and a total compensating error of around 55 W/m^2 suggesting a 35-W/m^2 error that is offset by a 15-W/m^2 error. While compensating errors play a reduced role over the Southern Ocean compared to the global data, it is clear that they are still important for understanding radiation biases within this region.

4. Conclusion

This work describes a set of approaches that can be used to compare model runs with differently tuned parameterizations. The average behavior of key variables are compared for each of the cloud clusters showing only minor differences between the model runs. Many of the clusters were identified as having distinct radiative biases over the Southern Ocean. Additionally, several clusters, such as clusters 5 and 6, were identified as being particularly impacted by the changes. The geographic distribution of SW CRE over the Southern Ocean was also examined, demonstrating that there are two distinct regional biases within the Southern Ocean and that improvements in one of these regions corresponds to downgrades in the other. A potential explanation for this relationship is that modifying the cloud parameterization leads to changes in the brightness of clouds that uniformly impacts the wider Southern Ocean region.

To better evaluate the accuracy of each of the model runs, a new approach was introduced that factors in both the mean bias and the bias without compensating errors. Overall these results indicate that the model runs with the smallest global biases had mixed results over the Southern Ocean. The model runs that altered capacitance were shown to reduce the magnitude of their compensating errors when combined with an additional set of changes to ice nucleation. Compensating errors were shown to play a larger role in the global radiative bias than in the Southern Ocean bias. This technique could also potentially be applied to analyzing different regions or to the comparison of other variables beside CRE. The method of breaking up data sets into clusters and then examining the errors in specific clusters to identify compensating errors could also be used much more widely in model evaluation efforts.

Acknowledgments
The authors would like to thank the teams behind MODIS, CERES, the HadGEM model, and the Deep South National Science Challenge. Funding for this project was provided as part of the Deep South National Science Challenge Cloud and Aerosols project. The CERES data were obtained from https://ceres.larc.nasa.gov/, and the MODIS data were obtained from https://ladsweb.modaps.eosdis.nasa.gov/. The HadGEM 3 model was created and is maintained by the U.K. Met Office. The data outputs used in this paper are accessible at the address https://doi.org/10.5281/zenodo.2230584.

References

Bodas-Salcedo, A., Andrews, T., Karmalkar, A. V., & Ringer, M. A. (2016). Cloud liquid water path and radiative feedbacks over the Southern Ocean. *Geophysical Research Letters*, *43*, 10,938–10,946. https://doi.org/10.1002/2016GL070770

Bodas-Salcedo, A., Webb, M. J., Bony, S., Chepfer, H., Dufresne, J.-L., Klein, S. A., et al. (2011). COSP: Satellite simulation software for model assessment. *Bulletin of the American Meteorological Society*, *92*(8), 1023–1043. https://doi.org/10.1175/2011BAMS2856.1

Bodas-Salcedo, A., Williams, K. D., Field, P. R., & Lock, A. P. (2012). The surface downwelling solar radiation surplus over the Southern Ocean in the Met Office model: The role of midlatitude cyclone clouds. *Journal of Climate*, *25*(21), 7467–7486. https://doi.org/10.1175/JCLI-D-11-00702.1

Bodas-Salcedo, A., Williams, K. D., Ringer, M. A., Beau, I., Cole, J. N. S., Dufresne, J. L., et al. (2014). Origins of the solar radiation biases over the Southern Ocean in CFMIP2 models. *Journal of Climate*, *27*(1), 41–56. https://doi.org/10.1175/JCLI-D-13-00169.1

Brown, A., Milton, S., Cullen, M., Golding, B., Mitchell, J., & Shelly, A. (2012). Unified modeling and prediction of weather and climate: A 25-year journey. *Bulletin of the American Meteorological Society*, *93*(12), 1865–1877. https://doi.org/10.1175/BAMS-D-12-00018.1

Furtado, K., & Field, P. (2017). The role of ice microphysics parametrizations in determining the prevalence of supercooled liquid water in high-resolution simulations of a Southern Ocean midlatitude cyclone. *Journal of the Atmospheric Sciences*, *74*(6), 2001–2021. https://doi.org/10.1175/JAS-D-16-0165.1

Furtado, K., Field, P. R., Boutle, I. A., Morcrette, C. J., & Wilkinson, J. M. (2016). A physically based subgrid parameterization for the production and maintenance of mixed-phase clouds in a general circulation model. *Journal of the Atmospheric Sciences*, *73*(1), 279–291. https://doi.org/10.1175/JAS-D-15-0021.1

Furtado, K., Field, P. R., Cotton, R., & Baran, A. J. (2015). The sensitivity of simulated high clouds to ice crystal fall speed, shape and size distribution. *Quarterly Journal of the Royal Meteorological Society*, *141*(690), 1546–1559. https://doi.org/10.1002/qj.2457

Gates, W. L., Boyle, J. S., Covey, C., Dease, C. G., Doutriaux, C. M., Drach, R. S., et al. (1999). An overview of the results of the Atmospheric Model Intercomparison Project (AMIP I). *Bulletin of the American Meteorological Society*, *80*(1), 29–55. https://doi.org/10.1175/1520-0477(1999)080<0029:AOOTRO>2.0.CO;2

Heymsfield, A. J., & Miloshevich, L. M. (2003). Parameterizations for the cross-sectional area and extinction of cirrus and stratiform ice cloud particles. *Journal of the Atmospheric Sciences*, *60*(7), 936–956. https://doi.org/10.1175/1520-0469(2003)060<0936:PFTCSA>2.0.CO;2

Hyder, P., Edwards, J. M., Allan, R. P., Hewitt, H. T., Bracegirdle, T. J., Gregory, J. M., et al. (2018). Critical Southern Ocean climate model biases traced to atmospheric model cloud errors. *Nature Communications*, *9*(1), 3625. https://doi.org/10.1038/s41467-018-05634-2

Jakob, C. (2003). An improved strategy for the evaluation of cloud parameterizations in GCMS. *Bulletin of the American Meteorological Society*, *84*(10), 1387–1402. https://doi.org/10.1175/BAMS-84-10-1387

Jin, D., Oreopoulos, L., & Lee, D. (2016). Regime-based evaluation of cloudiness in CMIP5 models. *Climate Dynamics*, *48*, 1–24. https://doi.org/10.1007/s00382-016-3064-0

Jin, D., Oreopoulos, L., & Lee, D. (2017). Simplified ISCCP cloud regimes for evaluating cloudiness in CMIP5 models. *Climate Dynamics*, *48*(1–2), 113–130. https://doi.org/10.1007/s00382-016-3107-6

Kay, J. E., Hillman, B. R., Klein, S. A., Zhang, Y., Medeiros, B., Pincus, R., et al. (2012). Exposing global cloud biases in the Community Atmosphere Model (CAM) using satellite observations and their corresponding instrument simulators. *Journal of Climate*, *25*(15), 5190–5207. https://doi.org/10.1175/JCLI-D-11-00469.1

Kay, J. E., Wall, C., Yettella, V., Medeiros, B., Hannay, C., Caldwell, P., & Bitz, C. (2016). Global climate impacts of fixing the southern ocean shortwave radiation bias in the Community Earth System Model (CESM). *Journal of Climate*, *29*(12), 4617–4636. https://doi.org/10.1175/JCLI-D-15-0358.1

Kohonen, T. (1998). The self-organizing map. *Neurocomputing*, *21*(1–3), 1–6. https://doi.org/10.1016/S0925-2312(98)00030-7

Kohonen, T. (2013). Essentials of the self-organizing map. *Neural Networks*, *37*, 52–65. https://doi.org/10.1016/j.neunet.2012.09.018

Mason, S., Fletcher, J. K., Haynes, J. M., Franklin, C., Protat, A., & Jakob, C. (2015). A hybrid cloud regime methodology used to evaluate Southern Ocean cloud and shortwave radiation errors in ACCESS. *Journal of Climate*, *28*(15), 6001–6018. https://doi.org/10.1175/JCLI-D-14-00846.1

McDonald, A. J., Cassano, J. J., Jolly, B., Parsons, S., & Schuddeboom, A. (2016). An automated satellite cloud classification scheme using self-organizing maps: Alternative ISCCP weather states. *Journal of Geophysical Research: Atmospheres*, *121*, 13,009–13,030. https://doi.org/10.1002/2016JD025199

Platnick, S., King, M. D., Ackerman, S. A., Menzel, W. P., Baum, B. A., Riédi, J. C., & Frey, R. A. (2003). The MODIS cloud products: Algorithms and examples from Terra. *IEEE Transactions on Geoscience and Remote Sensing*, *41*(2 PART 1), 459–472. https://doi.org/10.1109/TGRS.2002.808301

Platnick, S., Meyer, K. G., King, M. D., Wind, G., Amarasinghe, N., Marchant, B., et al. (2017). The MODIS cloud optical and microphysical products: Collection 6 updates and examples from Terra and Aqua. *IEEE Transactions on Geoscience and Remote Sensing*, *55*(1), 502–525. https://doi.org/10.1109/TGRS.2016.2610522

Schuddeboom, A., McDonald, A. J., Morgenstern, O., Harvey, M., & Parsons, S. (2018). Regional regime-based evaluation of present-day general circulation model cloud simulations using self-organizing maps. *Journal of Geophysical Research: Atmospheres*, *123*, 4259–4272. https://doi.org/10.1002/2017JD028196

Tan, J., Oreopoulos, L., Jakob, C., & Jin, D. (2017). Evaluating rainfall errors in global climate models through cloud regimes. *Climate Dynamics*, *0*(0), 1–14. https://doi.org/10.1007/s00382-017-3806-7

Walters, D., Baran, A., Boutle, I., Brooks, M., Earnshaw, P., Edwards, J., et al. (2017). The Met Office Unified Model Global Atmosphere 7.0/7.1 and JULES Global Land 7.0 configurations. *Geoscientific Model Development Discussions*, 1–78. https://doi.org/10.5194/gmd-2017-291

Wielicki, B. A., Barkstrom, B. R., Harrison, E. F., Lee, R. B., Smith, G. L., & Cooper, J. E. (1996). Clouds and the Earth's Radiant Energy System (CERES): An Earth observing system experiment. *Bulletin of the American Meteorological Society*, *77*(5), 853–868. https://doi.org/10.1175/1520-0477(1996)077<0853:CATERE>2.0.CO;2

Wild, M., Ohmura, A., Gilgen, H., & Roeckner, E. (1995). Validation of general circulation model radiative fluxes using surface observations. *Journal of Climate*, *8*(5), 1309–1324. https://doi.org/10.1175/1520-0442(1995)008<1309:VOGCMR>2.0.CO;2

Williams, K. D., & Tselioudis, G. (2007). GCM intercomparison of global cloud regimes: Present-day evaluation and climate change response. *Climate Dynamics*, *29*(2–3), 231–250. https://doi.org/10.1007/s00382-007-0232-2

Williams, K. D., & Webb, M. J. (2009). A quantitative performance assessment of cloud regimes in climate models. *Climate Dynamics*, *33*(1), 141–157. https://doi.org/10.1007/s00382-008-0443-1

Wood, N., Staniforth, A., White, A., Allen, T., Diamantakis, M., Gross, M., et al. (2014). An inherently mass-conserving semi-implicit semi-Lagrangian discretization of the deep-atmosphere global non-hydrostatic equations. *Quarterly Journal of the Royal Meteorological Society*, *140*(682), 1505–1520. https://doi.org/10.1002/qj.2235